(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 January 2001 (18.01.2001)

PCT

(10) International Publication Number
WO 01/04787 A2

(51) International Patent Classification[7]: G06F 17/00

(21) International Application Number: PCT/US00/19153

(22) International Filing Date: 12 July 2000 (12.07.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/143,617 13 July 1999 (13.07.1999) US
09/516,237 1 March 2000 (01.03.2000) US

(71) Applicant: ALLADVANTAGE.COM, INC. [US/US]; 4010 Point Eden Way, Hayward, CA 94545 (US).

(72) Inventors: O'CONNOR, Shawn, M.; 2205 Birdgepointe Parkway R326, San Mateo, CA 94404 (US). CROMWELL, Raymond, J.; 2205 Bridgepoint Parkway R326, San Mateo, CA 94404 (US).

(74) Agents: MALLIE, Michael, J. et al.; Blakely, Sokoloff, Taylor & Zafman LLP, 12400 Wilshire Boulevard, 7th Floor, Los Angeles, CA 90025 (US).

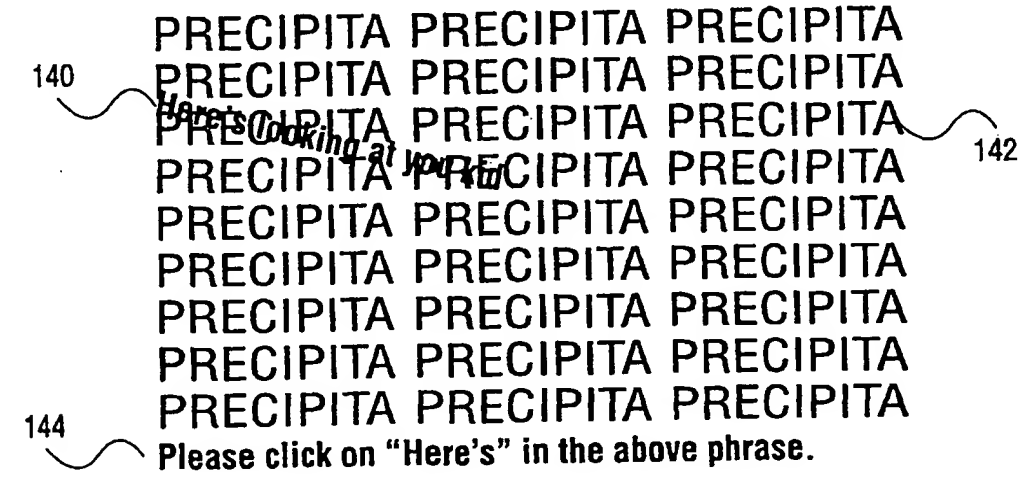
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD AND SYSTEM FOR CLASSIFYING USERS OF AN ELECTRONIC NETWORK



(57) Abstract: A system and method of classifying users of an electronic network to prevent receipt of unsolicited bulk or commercial email. The system includes a mail server for receiving email from a user. The mail server determines if the received email is from a source previously selected by the user. If the email sender is not in a database file of acceptable sources of email, a test is performed to test the humanity of the source of the email. If the source is classified as of human origin the email is presented to the user.

WO 01/04787 A2

METHOD AND SYSTEM FOR CLASSIFYING USERS OF AN ELECTRONIC NETWORK

This application claims the benefit of U.S. Provisional Patent Applications Serial No. 60/122,207, filed March 1, 1999, and Serial No. 60/143,617 filed July 13, 1999, the entire disclosures of which are incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates generally to systems and methods for sending and receiving electronic mail messages over an electronic network, such as the Internet, and, more specifically, to improved methods and systems for identifying and classifying the users of said electronic network. Moreover, a method is described by which certain sub-classes of users can be either 1) pre-classified by the system prior to the interaction (pre-qualification), or 2) exempted from the classification method (opted-out), when desirable. However, it is clear upon inspection of this specification, the claims and the drawings attached hereto, that the present invention may find application to a wide variety of circumstances involving human interaction in which it is beneficial to distinguish human interaction from non-human interaction, or to distinguish among two or more classes of human interaction.

BACKGROUND OF THE INVENTION

The increased use of electronic communication media such as electronic messaging (i.e., "email"), including the rapidly growing number of Internet users and acceptance of the Internet as a widely-available means of communication, has been accompanied by the advent of users and techniques that exploit, among other aspects, the medium's ability to very quickly disseminate information to an extremely large number of other users at little or no cost to the sender. In many situations, this capability results in many users receiving varying quantities of unsolicited and unwanted electronic mail messages, sometimes referred to as "spam." Such unwanted email has a number of detrimental effects, including, but not limited to, the subsequent time required for the recipient user to manually sort through a potentially large number of such email messages.

Further, the elements of a communication system, including, but not limited to, an electronic communication system, comprise a valuable economic resource. The presence of unsolicited email messages or spam propagating throughout an electronic network wastefully consumes the network's resources. This wasteful consumption has a particularly negative effect upon the traffic-sensitive aspects of a communications network; in particular, the network's communications and computing resources such as, but not limited to, servers, switches, access lines (e.g., T1 carriers), and routers within the network. The effect of the resulting waste therefore is to impose costs upon all users of an electronic communications network (except for the sender) as well as the network service provider.

In particular, Internet email has a number of features that make it attractive to those who would like to take undue advantage of such a widely deployed communication medium. One such feature is that most of the

aforementioned costs associated with sending bulk unsolicited bulk email (UBE) aren't incurred by the sender. Senders of UBE (i.e., "spammers") often shift the costs of delivering very large numbers of email to third-party mail transfer agents (MTA) that in turn incur the costs of bandwidth, message queuing, etc. Unfortunately, the economics are such that UBE will continue to be a problem unless a cost can be imposed on senders of millions of UBE messages. The spam problem persists because of how email costs are distributed, leading to, among other problems, a tragedy of the commons scenario that affects all users of an electronic network.

One proposed solution to this problem is to impose an email postage cost upon the sender or spammer. In other words, deter UBE by charging a nominal amount of postage for each email sent. Bringing the costs of sending email to even a fraction of a cent per message sent would likely dramatically reduce present abuses, since senders of UBE (spammers) would then have a financial burden to overcome. The problem with such micropayment postage approaches is that the installed base of all email clients and servers would have to be modified to support use of any such system. Such an undertaking would be cost prohibitive. Attempts at online micropayment systems have met with significant deployment friction, since deploying such a system has to overcome the understandable fear of breaking a functioning and mission-critical system relied on by millions of people. Further, this technique might run counter to cultural norms established within the Internet community respecting free email transport.

Other methods that rely on sending a password or having senders email back a message in response to an automatic reply can be used to block email without a valid return address (a common feature of UBE). However the simple

text nature of these sorts of messages make it very possible to automate the process of replying to automatic confirmation email. Further, these systems often block desirable but automated email sent by agents simply not programmed to handle the auto-replies.

Other attempts to filter spam based on email headers, message content, source addresses, etc. only solve parts of the problem and dedicated spammers can and have routinely bypassed these filters.

Alan Turing proposed a test for 'human level' intelligence called the Turing Test, which is well-known in the art, in which an interview is conducted through a text terminal. If the interviewer is unable to determine through questioning whether the interviewee is a human being, then the interviewee would be designated as possessing human level intelligence. Further details regarding the Turing test may be found in Turing, Alan M., Computing Machinery and Intelligence, *Mind, 59*, 433-560 (1950), the teachings of which are herein incorporated by reference. This technique, however, has not been used to distinguish and classify users of an electronic communications network in the manner described herein, and even if so employed, the traditional (interview) Turing test would take much longer to produce more ambiguous results.

SUMMARY OF THE INVENTION

Therefore, it is a general object of the present invention to provide a method and system that distinguishes between human and non-human users of an electronic communications network.

A further general object of the present invention is to provide a method and system that reduces the costs imposed upon recipients of unsolicited bulk

email (UBE) in terms of the recipient's time required to respond and process such email, as well as consumption of computing and communications resources.

A still further general object of the present invention is to provide a method and system that can impose costs on senders of bulk email in order to discourage abuses of network resources or the time and attention of intended recipients.

A still further object of the present invention is to provide a method by which solicited or otherwise desired yet automated email can be identified and allowed to reach recipients who are utilizing the present invention to block automated incoming email.

A still further general object of the present invention is to provide a method and system that accomplishes these objectives while providing high immunity from countermeasures likely to be employed in attempts to thwart or overcome its beneficial effects.

The invention is a system and method of classifying users of an electronic network to prevent receipt of unsolicited bulk or commercial email. The system includes a mail server for receiving email from a user. The mail server determines if the received email is from a source previously selected by the user. If the email sender is not in a database file of acceptable sources of email, a test is performed to test the humanity of the source of the email. If the source is classified as of human origin the email is presented to the user.

These as well as other objects are apparent from inspection of this specification and the drawings attached hereto.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a functional block diagram of a user classification system;

Figure 1A shows the presentation to a user of a preferred embodiment of an attention/humanity test;

Figure 2 is an illustration the user classification system applied to the UBE problem;

Figure 3 is a sequential flow diagram of the user classification system applied to the UBE problem;

Figure 4 is an illustration of the user classification system applied to an online form submission context;

Figure 5 is an illustration of the user classification system applied to an online web link traversal control application;

Figure 6 is a functional block diagram of a preferred embodiment of the user classification system;

Figure 7 is a functional block diagram illustrating the operation of the user classification system;

Figure 8 is a state flow diagram of the user classification system;

Figure 9 describes generation of the test image URL;

Figure 10 describes decoding and presentation of the test image to the sender;

Figure 11 describes test results checking;

Figure 12 describes generation of limited use email addresses that can be given to solicited or otherwise desired automated mailers; and

Figure 13 describes the checking of a limited use email address when used.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention relates to systems and methods for quickly identifying and classifying users of an information network or electronic network,

such as the Internet, by reliably distinguishing between human beings and computer programs in various online contexts. In a presently preferred embodiment, a user classification system is described for an exemplary application to block receipt of unsolicited bulk email (UBE), or junk email or spam. However, it is apparent upon inspection of this specification, claims and the drawings attached hereto, that the user classification system may find application to a wide variety of situations or alternative embodiments in which it is useful to distinguish a human being from an automated process such as a computer program. The user classification system can, without limitation, quickly classify, indicate, or flag as suspect any interaction with an electronic network or information network that isn't originated by a human. Examples of applications in which the methods and systems disclosed herein are useful may include, but are not limited to: distinguishing unsolicited bulk email (i.e., spam) sent by an automated process from unsolicited email sent by a human being; distinguishing human beings from robots clicking on links online (e.g., website advertisers paying by the click don't want to pay for synthetic hits from search engine indexing robots or from web sites using robots to artificially increase their hit-counts); and, distinguishing astroturfing in poll and survey results (i.e., the generation of automated poll respondents or responses, or spamming a survey, such that the poll or survey results may be improperly skewed).

The user classification system imposes a cost in terms of human attention; however, unlike the proposed cash or near-cash instruments of micropayment postage schemes, the user classification system works with the installed base of network email infrastructure without requiring extensive modification of the entire installed base.

In a cognitive sense, a human can pay (or, pay with) attention with low marginal cost to the human; computers, however, are not yet capable of the same high-level cognitive functions that come naturally to humans. Therefore, computers (or automata) must incur a relatively high marginal cost to approach some of the cognitive capabilities of humans, if it is possible at all.

Therefore, rather than employing monetary postage to alter the economics of sending UBE, the user classification system addresses the method of using automated means, such as a computer program, to send such bulk email. Specifically, in a presently preferred embodiment, the user classification system filters out email sent programmatically while allowing email sent by human beings to be received by a user.

Conducting an interview would be one method to determine interaction with a human intelligence, but the costs of doing so might be more than simply enduring spam. The user classification system therefore utilizes an attention/humanity test that presently relies on human ability to quickly process various sensory input, including, but not limited to, complex visual images. Examples of the attention/humanity test include, but are not limited to, putting up an image of, for example, a dog, a cat, and a bear with the challenge "click on the picture of a dog"; or to make this example even more challenging for a machine, "click on the animal least likely to be a household pet." Such activities are very easy for a sighted human being, but extremely difficult for current programmed computers. Humans can pay (with) attention, computers can't.

Spammers often obtain email addresses by harvesting mailing lists and other online forums, discussion groups, usenet news, etc. After posting messages in public fora, people soon learn how this works. One technique in use to counter the programs scanning for email addresses is to corrupt (or 'munge')

the email address used in posting. For example, for the email address "user@location.com" one possible anti-spam address for corresponding usenet posts might look something like "usernospam@nospamlocation.nospam" with perhaps a note somewhere in the text of their messages containing instructions on how to recreate the user's true email address by eliminating all the "nospam" from the address and appending a ".com" or a similar transformation that would result in the correct email address. This technique is essentially a method to alter the address in a way intended to be difficult for programs to correct but relatively easy for humans to fix. However, this munging method has limitations, including:

1. Though it is relatively easy for a human to figure out and hard for a program to fix email addresses treated in this way, the munging method breaks the way user software (e.g., mail, news client software) works. In other words, users expect, and their software is configured, to be able to send email where your headers, etc. say you can receive email. Munging renders the publicized email address unusable. It is cumbersome to have to locate and refer back to the original message with the instructions on how to reconstruct an email address so that it works after initially sending email and only later getting a delivery error. In contrast, the present invention works with existing email systems by filtering-by-sender-type upon receipt of an email, rather than when publicizing an email address.

2. A further problem with the munging method is that the address need only be reconstructed once before it can suffer from UBE. Senders of UBE might find it cost effective to fix by human means each of these addresses once in order to send email to it countless times or sell the corrected address in mailing lists, etc. Theoretically, the munging technique only imposes a cost once after which the fixed email can be shared with other (even UBE) senders. In contrast, the

present invention imposes a cost each time a new sender sends an email to a recipient, thus allowing users to use and publicize a working email address all they like. The classification system works by presenting a test in order to determine whether senders of email (in this case) are human beings or automata. In a presently preferred embodiment, the user classification system presents an attention/humanity test that is very easy for humans to pass but difficult for a program to pass.

The user classification system quickly distinguishes humans (and human action or interaction) from automated actions (e.g., generated by machines, computers, or programs) in order to treat them differently if so desired. In an exemplary application of a presently preferred embodiment, the user classification system filters UBE, sent by automaton, by filtering out those messages originating from a sender that has failed a test that is very easy for a human paying attention to pass.

Figure 1 illustrates use of the abstract attention/humanity test used to quickly distinguish human interaction from interaction with automata such as computer programs. An action shown as block 120, starts the process. A test attention/humanity check is made in step 122. If pass, then the response was likely human , block 124, if fail, then is unlikely to be human, block 126.

In a presently preferred embodiment, the attention/humanity test appears to the user as indicated in Figure 1A. The test image 140 has text 142 in the background and employs several means of making it difficult to defeat programmatically by, for example, scanning the test via optical character recognition (OCR). As such, the attention/humanity test requires not just cognition or reading of written text, but also requires visual perception and cognition of spatial orientation. An example question is shown as 144.

In a presently preferred embodiment, the attention/humanity test therefore requires the user to possess textual, visual, and spatial cognition or awareness, the ability to integrate this awareness of multiple such relationships in context and the ability to determine and act on an appropriate response (e.g., selecting the correct responsive choice or otherwise responding appropriately to the given challenge). The present embodiment of the test is dynamically generated in order to present a different test for each instance and alternative embodiments of the attention/humanity test will be made even more difficult to pass programmatically. Alternative embodiments may also require a user to be cognizant of multiple sensory inputs such as, but not limited to, written text, spoken text, verse, song, lyrics, or other such auditory or visual indications or responses. It is to be recognized that many alternative embodiments employing other words, phrases, arrangements of letters, icons, more complex or photographic images, as well as multimedia, auditory, or visual indications are possible within the spirit and scope of the present invention. Further, the user classification system makes the human/non-human determination quickly, whereas a traditional Turing test could take several minutes or hours before a determination is made.

Figure 2 illustrates a presently preferred embodiment of the user classification system in which the attention/humanity test is applied to the UBE problem. A sender sends an email to a recipient, step 200. In the process of handling incoming email from various senders to various recipients, the user classification system determines whether an attention/humanity test needs to be administered for this (sender, recipient) pair, step 204. If the user classification system determines that an attention/humanity test is required, the user classification system administers the test, again in step 204, in order to determine

whether the incoming email is likely to be bulk/automated UBE. If the sender passes the attention/humanity test, then the email is delivered to the recipient user as normal email, block 206. If not, the email message is treated as UBE, block 208. In a presently preferred embodiment, user classification system 100 will classify or sort UBE such that it is treated differently from non-UBE; this may include, but is not limited to, providing UBE messages in a separate UBE file folder in the recipient's web browser interface. Other means of distinguishing UBE to the recipient are within the spirit and scope of the present invention. In an alternative embodiment, user classification system 100 uses the determination of UBE as a criterion for ordering or prioritizing a single, integrated list of unread email delivered to the recipient user instead providing UBE in a separate file. More specifically, unread email from individuals to other individuals or to a small group may be accorded relatively greater priority than unread UBE, the relative priorities of unread email messages being indicated by order of appearance in the user interface. Alternatively, various means of indicating this relative prioritization of unread email may be provided, including captioning, color coding, etc.

Figure 3 is a sequential flow diagram of a presently preferred embodiment of the user classification system in which the attention/humanity test is applied to the UBE problem. The user classification system maintains a list of senders who have already taken and passed the attention/humanity test so that senders are only asked to pass the test once (or some recipient-determined number) per person emailed. Further, if a particular sender accrues too many failed attempts, the user classification system determines an error condition and will stop presenting tests, leaving the associated email to be treated as not having passed the test. In a presently preferred embodiment, this is accomplished by setting a

numerical limit on the number of attention/humanity tests sent in response to a particular incoming email over a period of time. This is a protection against repeated attempts to overcome the attention/humanity test of the user classification system (e.g., statistical or other attempts to pass the test programmatically) in which a program sender is likely. In an alternative embodiment, the user classification system provides means for the user to set the limiting parameters regarding when to cease sending attention/humanity tests to a particular sender. Such user-determined parameters may include, but are not limited to: allowing all emails to be received; only allowing receipt of emails for which the sender has passed multiple tests, wherein the number of tests is controlled by the user; only allowing receipt of emails from particularly identified senders or groups of senders; or any combination of these as well as other limitation means. It is apparent that a variety of limitation means may be employed within the spirit and scope of the present invention in a manner that imposes a relatively small burden on the email recipient user.

The following further illustrates the sequential process flow of a preferred embodiment of the user classification system according to Figure 3. An originating user (sender) sends email, step 300, to an intended recipient using the user classification system, 302, to block UBE. The recipient's email server (Mail Transfer Agent, or MTA), 304, receives the sender's email and checks to see if the sender should pass a test prior to delivery. There are a few instances in which a sender may not need to take a test and the email is delivered in step 306. These situations include:

1. The recipient has already approved delivery of email from the sender, perhaps delivery under certain usage conditions and the current email falls within those usage limits (e.g., until further notice pass it through, pass

through only four per month, 12 per year, etc. determined by the user's approval and usage settings).

2. The email is being sent to the recipient from a sender that took the test before sending an email (e.g., a user who reads and responds to email offline, logs on briefly only to deliver/receive email, then logs off may not receive test messages triggered by his outgoing email until his next login, such a user might opt to take tests first then email in order to avoid the possibility of delaying his message delivery until his next online session).

3. An automated agent is sending to a pre-approved limited use address. The construction and verification of these addresses is described in Figures 12 and 13. These sorts of pre-authorized limited-use addresses can be used to allow desirable automated mailings such as mailing lists or electronic greeting cards or opt-in (solicited) commercial email sent by programs (such as order confirmations, newsletter registrations, etc).

4. In an alternative embodiment, if the sender is identified by the user classification system as a "trusted" sender of automated email (i.e., is listed in the reputation/trustability database).

If the determination is made to administer the test is made in step 304 for an incoming email (e.g., by inspection of sender/recipient, etc.), then the incoming email is queued and a test URL is generated and sent to the sender's email address, step 308. The sender upon receiving the test, step 312, would then take the test, step 314, in his HTML-enabled email client (mail user agent, or MUA) or open the test URL in a web browser. If the sender passes the test then the queued message related to that particular test is treated accordingly (e.g., delivered to the intended recipient's mailbox, etc.), step 316.

If the test is failed then another can be generated on the fly just in case it

was user error, mistake, etc. After a few tries, an error message, step 318, will indicate that no more attempts will be allowed and the result of the test is that the sender is very likely a program. Likewise, never taking the test is treated as a failed test result (messages that require a test begin and remain in the test-failure state until the test is passed).

Figure 4 illustrates a presently preferred embodiment of the user classification system applied to an online form submission context. A user submits a form in step 400. The user classification system intercepts the user's submission of the form, step 402, and provides the attention/humanity test, step 404, to the submitting user to ascertain whether the submitting user is a human or automata (e.g., computer program). Non-human submissions may be blocked from further reception by the receiver, step 406, in order to exclude programmed form submissions designed, for example, but not limited to, skew the results of a poll or otherwise frustrate or mislead the recipient's goals. Submissions passing the test are accepted in step 408.

Figure 5 illustrates a presently preferred embodiment of the user classification system applied to an online web link traversal control. A user clicks on a web link in step 500. The user classification system intercepts the user's sequential web link selections, in step 502, or clickthroughs, and provides the attention/humanity test to the user at one or more points of the link traversals, step 504, to ascertain whether the user is a human or automata (e.g., computer program). Non-human users may be blocked from further link traversal, step 506, or otherwise treated accordingly in order to protect against programmed modes of attack or frustration, such as, but not limited to, indexing by web robots or programmatic inflation of usage statistics. User clicks deemed valid are passed through in step 508. Alternatively, the user classification

system may be used in a similar manner by online advertisers to distinguish actual human user selections, or clicks, from synthetic user clicks that may otherwise inflate usage statistics. Alternatively, the user classification system may be used in a similar manner to distinguish human users from so-called non-idle programmed users, for example an Internet Service Provider (ISP) inquiring (usually with a dialog box) if an idle human user wishes to remain online. A common practice of ISP customers is to frustrate the intent of the ISP by running simple programs that always answer yes to the question of remaining online, even if they are truly away. ISPs could use an alternative embodiment of the present invention in order foil the programs by which users automate their responses to these inquiries, and so more accurately determine which of their users are truly idle or away, etc. The user classification system is able to distinguish between human users that are paying attention and human users who are not paying attention, even though they may be employing an automated agent to claim they are paying attention.

Referring to Figure 6, in a presently preferred embodiment, a user classification system 100 comprises receiving mail server 101, a test server 102, a recipient 103, and a sender 104. Mail server 101 receives mail for recipient 103. Mail server 101 may be any mail transfer agent (MTA) present in an electronic network such as, but not limited to, the Internet or an intranet. Mail server 101 blocks UBE by checking incoming email to see if it originates from an address in recipients allowed list 106. In a presently preferred embodiment, recipients allowed list 106 is located at mail server 101. Mail server 101 interfaces to test server 102 via standard interfaces including, but not limited to a shared database, POP, IMAP, SMTP, and HTTP. The attention/humanity test is located at test server 102. When mail server 101 determines that an attention/humanity test is

to be administered as described herein, mail server 101 generates the URL of the test and sends an email to sender 104 (at the sender's email address) containing the test URL. Sender 104 then receives and views the test on test server 102. Sender 104 can then submit test results to test server 102 (via HTTP) at which time test server 102 determines whether sender 104 has passed or failed the test. Test server 102 will only give sender 104 a few chances to pass the test and if sender 104 fails them all the results will be recorded as failure and the corresponding email will be treated accordingly. Mail server 101 further comprises gateway software 105 and a recipients allowed list 106. Recipients allowed list 106 further comprises usage parameters. In an alternative embodiment, recipients allowed list 106 further comprises a reputation/trustability database indicating senders of desirable UBE. Recipient 103 comprises a receiving user, a computing device such as, but not limited to, a personal computer, wherein said personal computer further comprises standard peripherals including a modem, and client side software including, but not limited to, an email client. Sender 104 comprises a sending user, a computing device such as, but not limited to, a personal computer, wherein said personal computer further comprises standard peripherals including a modem, and client side software including, but not limited to, a web browser and an email client. Gateway software 105 is server side software that performs computations required to implement functions generally related to determining when an attention/humanity test is to be administered, as well as interpreting the test response/results, as described herein. In a presently preferred embodiment, gateway software 105 is implemented in the JAVA programming language, as are other software components comprising a presently preferred embodiment of user classification system 100. For further details regarding Internet message format

and protocols, see David Strom and Marshall T. Rose, *Internet Messaging*, Harcourt Brace; ISBN: 0139786104; Paperback - 400 pages (July 1998), the teachings of which are herein incorporated by reference.

Senders 104 can get on recipients allowed list 106 by either passing the attention/humanity test or through recipient 103 pre-approval. In an alternative embodiment, senders 104 are added or removed from the reputation/trustability database by a system administrator based on pre-defined criteria. Email from trusted senders 104 identified in the reputation/trustability database are indicated to recipient 103 as desirable UBE. If mail server 101 determines that an incoming email should be tested prior to delivery, sender 104 is sent the attention/humanity test in email. If sender 104 successfully passes the test, the original email is delivered to recipient 103. If sender 104 fails the test or fails to take the test, the original incoming email is treated accordingly (e.g., remains in a junk mail folder, or can be otherwise dealt with).

A further functional block diagram illustration of user classification system 100 is provided in Figure 7. In Figure 7 an email sender 700 uses an email client 702 to send email. A mail transfer agent 704 performs a test on the email sender through web client 706. If the sender is approved, a second email client 708 delivers the email to the intended recipient 710.

Figure 8 provides a state flow diagram of user classification system 100 for the case of a successfully completed attention/humanity test.

In a presently preferred embodiment, the attention/humanity test is designed principally to make it difficult for programs to pass but easy for humans, so if a sender passes the test it is very likely a human being, and if a sender fails the test it is very likely a machine. In an alternative embodiment, the attention/humanity test may be designed to distinguish between or among a

plurality of different classes of humans, for example, but not limited to, adults and children, attentive humans and inattentive humans, or telemarketers and non-telemarketers. Interactions can then be handled in the appropriate manner for each case. Recipients can determine what to do with email likely to originate from bulk email programs, whether to place it in a separate junk or low-priority mail folder, otherwise mark it as suspect, delete it, or just pass it through.

In a presently preferred embodiment, the attention/humanity test is delivered by a test server and can be accessed by any HTML enabled email viewer or web browser. The test image URL is generated as described in Figure 9 and is decoded and presented to the sender as specified in Figure 10. Shown in Figure 9, a random test parameters are chosen, step 900. Next in step 902 the parameters are packed into a string of bytes. The string is encrypted in step 904, and encoded as an integer into a URL, step 906.

In Figure 10 is shown the test generation. A test URL is decoded in step 1000. The parameters are decrypted in a string step 1002 and unpacked into the test parameters in step 1004. A check is made to see if a test has already been made previously in step 1006. A decision is made in step 1008, if yes then stop issue failure notice, step 1010. If no, draw background and noise as in step 1012. Using instructions coded in the parameters is form, step 1014. This is encoded as a GIF or JPG file in step 1016 and presented to the user in step 1018.

User classification system 100 checks the test results as specified in Figure 11. Shown is step 1100 wherein a user submits a test response. The URL is decoded to receive encrypted string and test results in step 1102. Strings are unpacked into test parameters in step 1104. A area for placement of a chosen word and surrounding "hotspot" is made in step 1106. Decision step 1108 determines is hotspot is clicked upon, yes block 1110 or no, block 1112.In a

presently preferred embodiment, the attention/humanity test is different each time presented because it is generated dynamically each time presented using certain parameters and certain random input in order to combat attempts to circumvent the user classification system. In a presently most preferred embodiment, test generation chooses test parameters from within a recipient-specified affinity group to allow for user customization and personal expression in test generation. For example, a recipient may choose test parameters that include material or information from a favorite television program or movie. A recipient is thereby able to use the attention/humanity test as a personalized method of introduction, online business card, or enjoyable puzzle or test that is provided to senders wishing to send email to the recipient.

In some circumstances, it is necessary to approve a proxy to deliver a message or to pre-approve a sender in such a way as to detect if that sender gives out your email address. One example of proxy approval is the case of online greeting cards. Online greeting cards are sent from the originating address of the company providing the service to the recipient, but they truly originate from the user who initiates the request. Accordingly, there needs to be a solution to allow an originating user to approve a third party to send email to a recipient. The first technique to accomplish this -- applicable to greeting cards -- is to allow an originating user to submit his email address, the email address or domain of the proxy, and the recipient email address to the user classification system, administer the attention/humanity test, and pre-approve the third party for a single one-time use. Figures 12 and 13 describe how a presently preferred embodiment of the user classification system accomplishes these objectives.

Further, there are some circumstances, however, where one needs to approve a whole range of originators who send messages via a proxy. Examples

of such situations include electronic discussion or mailing lists, wherein an originating user sends a message to the list itself, and an electronic mailing list service broadcasts this message to all of its recipients.

If users of the user classification system sign up to an electronic mailing list using their normal bulk-email-blocking email address, someone sending email to the mailing list could potentially get hundreds of attention/humanity tests after sending a single message (if that many recipients' mail servers determine that the attention/humanity should be sent). What is needed is a way to pre-approve messages originating from a mailing list, and only that mailing list. The present invention addresses this problem by providing a method to pre-approve email originating from the mailing list (and other desirable bulk mailers) by means of a unique, specially encoded, alternate email address for the recipient.

The aforementioned technique functions as digital pre-paid postage and is useful in a variety of contexts where users (later recipients) can give out alternate email addresses to senders encoded with usage limitations for the sender as well as the sender's email address or other identifying characteristics. This is accomplished as follows in accordance with Figure 12. First, the recipient enters sender address and user limits, step 1200. The sender to be pre-approved (a mailing list or other third party), is assigned a unique ID, step 1202. This unique ID is then combined with another number which represents options for this originator, step 1204, such as a limit on the number of pre-approved emails, a limit per unit time, an expiration date, etc. The numbers are simply packed together, bit-wise. Next in step 1206, they are encrypted with a standard encryption algorithm such as DES, IDEA, or Blowfish. Then in step 1208 they are separated into n-bit pieces, which can represent a number in the range of 0 to 2^n-1. These pieces are then used as an index in step 1210, for a table of 2^n words

and the corresponding words are then concatenated with '.' characters, and appended to the receiver's email address to form a human readable unique email address encoded with pre-approval and usage limitations between the sender and the receiver, step 1212.

An example of such an email address, generated for the receiver "johnsmith@precipita.com" in order to pre-approve email from mailinglist@location.com might be johnsmith-small.bike@precipita.com. Shown in Figure 13 is the use of the limited access email. In step 1300 mail is received for limited use address. Next a look up of integer pair for ward pair in address in made in step 1302. The 14 bit number pair is appended and decrypted to produce sender ID and use limit code in step 1304. A test is made in step 1306 to determine is the mailing is within bounds of usage limit code for the sender. If so, the email is delivered in step 1308, if not, a send failure notice is made, step 1310.

If the user does not wish a human readable address with the encoding, then the pre-approved email address may simply have the encoded ID and options concatenated as a number, such as "johnsmith-57637562@precipita.com."

After generating such an address, a user is free to subscribe to a mailing list (or other desirable bulk emailer) using it. Messages sent from the mailing list, and any user on the list, to the pre-approved email address will not be subject to a test. To prevent abuse of pre-approved addresses, the user classification system may automatically "renew" a user's subscriptions to internet mailing lists by unsubscribing, as well as revoking, the pre-approved nature of the old unique pre-approved email address, and resubscribing a newer unique pre-approved email address, on a regular basis. One method is for mailing lists to prevent

indiscriminate spamming by subjecting each subscriber's first post to moderation and only allowing subscribers to post to the list. For the case of pre-approval for a programmatically generated greeting card, etc., the pre-approval addresses may have a built in expiration mechanism, after which incoming email will no longer be pre-approved.

In an alternative embodiment, user classification system 100 allows email originating from one or more pre-approved senders of email to be received by the email recipient without an attention/humanity test first being successfully administered to the sender. In this embodiment, recipients allowed list 106 further comprises a reputation/trustability database in which one or more metrics are used to compute a threshold determination as to whether or not a particular email sender's email will not be treated as UBE as described herein. Thus, email sent by one of these "trusted" senders (i.e., a sender appearing in the reputation/trustability database) will be treated as UBE, but will be provided to the recipient user with an indication that the UBE comes from a trustworthy source -- even if the recipient user neglects to provide the trusted sender with a pre-approved email alias or proxy as described herein. In this embodiment, a system administrator will update and maintain the reputation/trustability database by adding or removing trusted senders according to a set of specified criteria, or by adjusting one or more metrics associated with the set of specified criteria, such as, but not limited to, the positive reputation of the sender, the value of the sender's information, the persistency of the sender's operations, and the quality of the sender's information. This approach is preferable to other methods which block email receipt from a sender based on a database of "blacklisted" senders, because such blacklisted "spammers" may change source address frequently or employ other means in order to frustrate effective use of

the blacklisting technique.

In support of this alternative embodiment, or independently, user classification system 100 may include a plurality of bulk email classifications. That is, instead of a binary determination of either human-originated email (i.e., individual to individual) or automatically-generated email (i.e., email sent programmatically with little regard to the identity of ultimate recipients, and which may or may not involve a human being present somewhere in the chain of transmission), user classification system 100 may also classify email to at least one intermediate classification, such as, but not limited to, desirable or trustworthy UBE.

A preferred embodiment of the present invention is implemented in source code using the JAVA (and PERL) programming language.

Thus, a method and system for classifying users of an electronic network has been shown that distinguishes between human and non-human users of an electronic communications network. Further, the present invention provides a method and system that increases the costs imposed upon senders of unsolicited bulk email (UBE) and thereby discourages abuse of email infrastructure and lowers the costs of using email for those utilizing the present invention (e.g., less time wasted dealing with spam) while providing high immunity from countermeasures. In particular, given that many resource utilization and network abuse problems can be ameliorated by imposing a usage cost, but also given that previously proposed solutions are either ineffective or involve significant upgrades of email infrastructure (and are therefore impractical), the present invention has the advantages of imposing a non-cash cost compatible with present infrastructure on certain actions, such as sending of bulk email or UBE. Further, the user classification system employs an acceptable non-cash cost

wherein the cost imposed is, essentially, attention of the sort which is not overly cumbersome for humans but relatively expensive for computer programs or automata. Specifically, cost is imposed in the form of a quick visual challenge-response test which is provided in a form difficult to subvert programmatically.

The user classification system described herein has been presented in a presently preferred embodiment for an application to unsolicited bulk email (UBE). However, it is clear upon inspection of this specification, and the appendices and drawings attached hereto, that the present invention is applicable to a wide variety of circumstances involving human interaction in which it is beneficial to distinguish human interaction from non-human interaction, or to distinguish among two or more classes of human interaction.

WHAT IS CLAIMED:

1. A system for classifying users of an electronic network comprising:

a mail server for receiving email from a sender;

a test server electrically interconnected with said mail server for testing the humanity of said received email at the request of said mail server;

a recipient for receiving email from said mail server;

said mail server checking email received to present only selected classes of email to said recipient.

2. The system of claim 1 wherein said mail server further contains a database of permitted senders.

3. The system of claim 1 wherein said mail server further contains a gateway for determining the need for humanity testing.

4. A method for classifying the users of an electronic network comprising the steps of:

receiving email from a sender at a mail server;

said mail server determining whether the source of the email is on an approved list;

testing the source of the email if not on said approved listing to classify if the sender is human; and

forwarding said email to recipient only if the source is classified as human.

5. A computer-readable media that causes a computer system to classify users of an electronic network by performing the steps comprising;

receiving email from a sender at a mail server;

said mail server determining whether the source of the email is on an approved list;

testing the source of the email if not on said approved listing to classify if the sender is human; and

forwarding said email to recipient only if the source is classified as human.

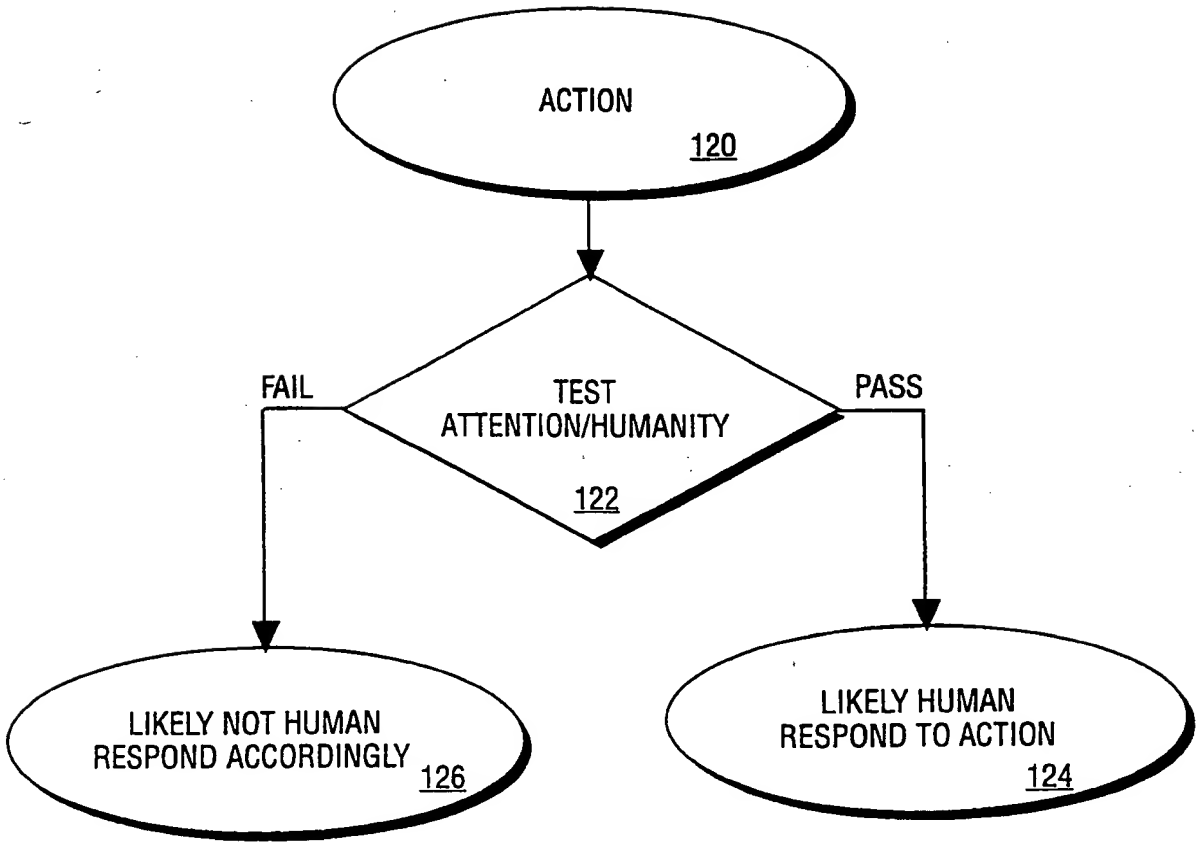

ACTION
120
FAIL
TEST
ATTENTION/HUMANITY
122
PASS
LIKELY NOT HUMAN
RESPOND ACCORDINGLY
126
LIKELY HUMAN
RESPOND TO ACTION
124

FIG. 1

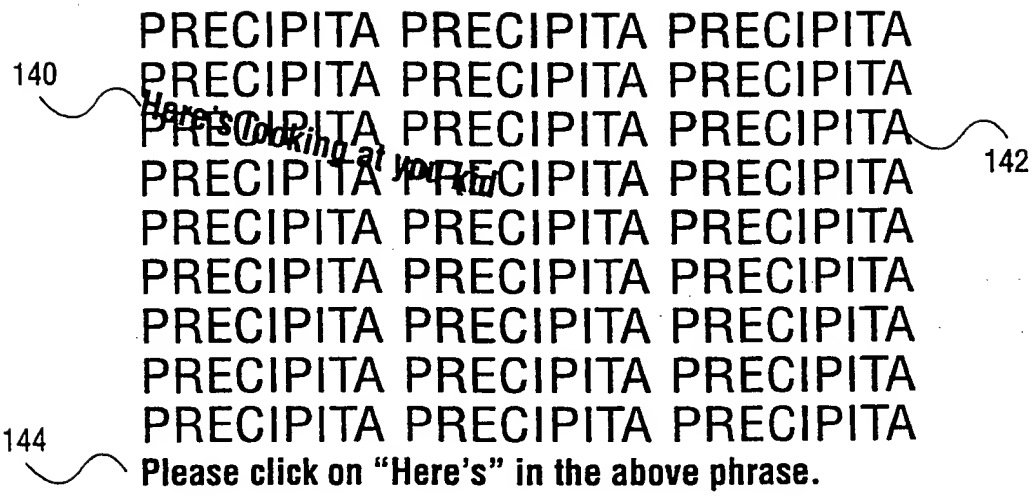

PRECIPITA PRECIPITA PRECIPITA
140
PRECIPITA PRECIPITA PRECIPITA
Here's looking at you, kid
142
PRECIPITA PRECIPITA PRECIPITA
PRECIPITA PRECIPITA PRECIPITA
PRECIPITA PRECIPITA PRECIPITA
PRECIPITA PRECIPITA PRECIPITA
PRECIPITA PRECIPITA PRECIPITA
144
Please click on "Here's" in the above phrase.

FIG. 1A

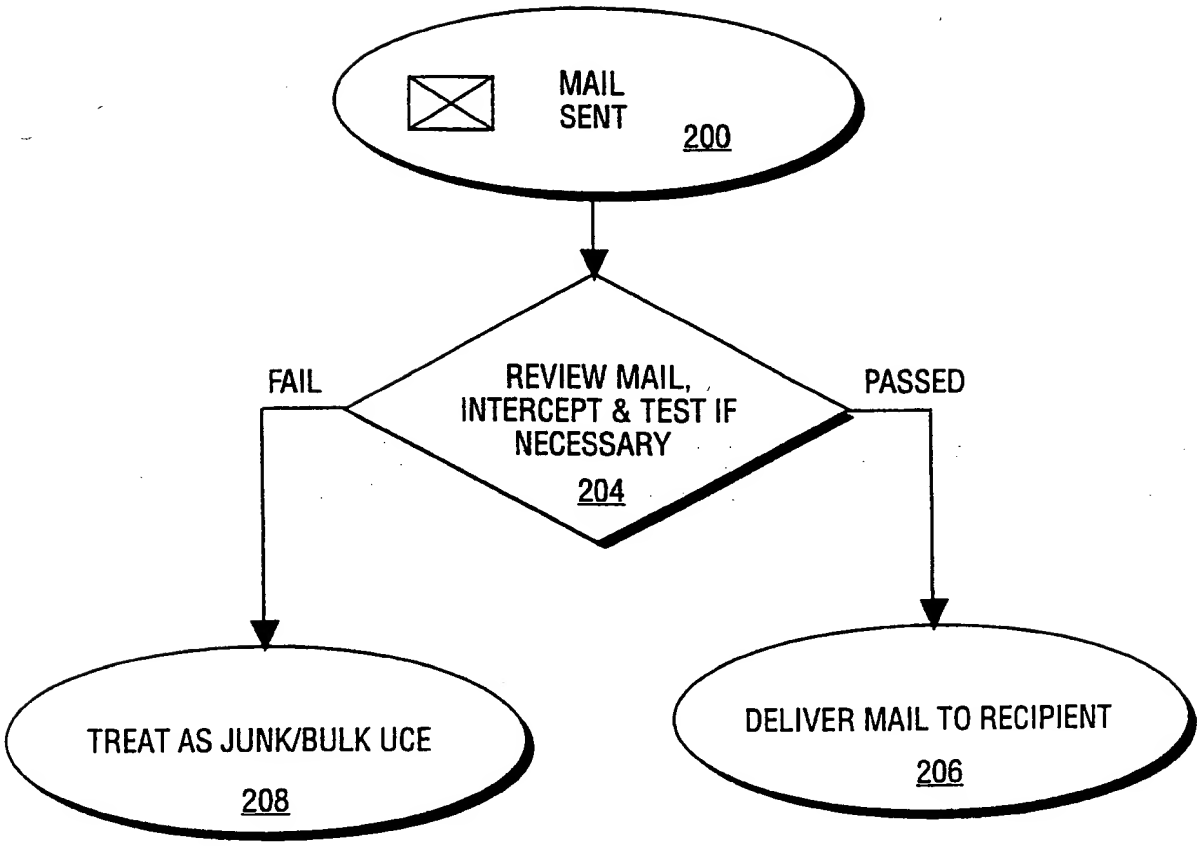

MAIL SENT
200
REVIEW MAIL, INTERCEPT & TEST IF NECESSARY
204
FAIL
PASSED
TREAT AS JUNK/BULK UCE
208
DELIVER MAIL TO RECIPIENT
206

FIG. 2

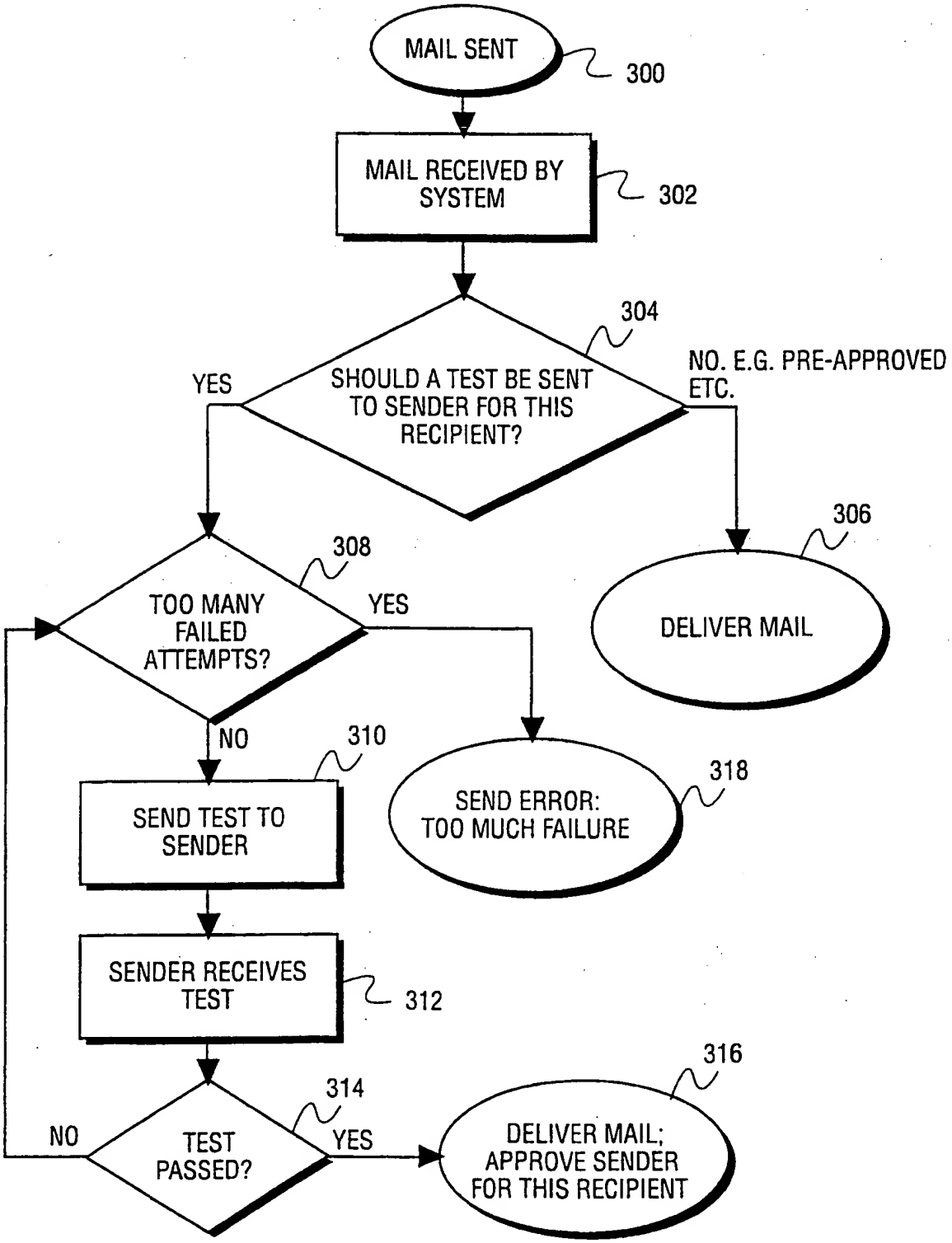

MAIL SENT
300
MAIL RECEIVED BY SYSTEM
302
304
SHOULD A TEST BE SENT TO SENDER FOR THIS RECIPIENT?
YES
NO. E.G. PRE-APPROVED ETC.
306
DELIVER MAIL
308
TOO MANY FAILED ATTEMPTS?
YES
NO
310
SEND TEST TO SENDER
318
SEND ERROR: TOO MUCH FAILURE
SENDER RECEIVES TEST
312
316
314
NO
TEST PASSED?
YES
DELIVER MAIL; APPROVE SENDER FOR THIS RECIPIENT

FIG. 3

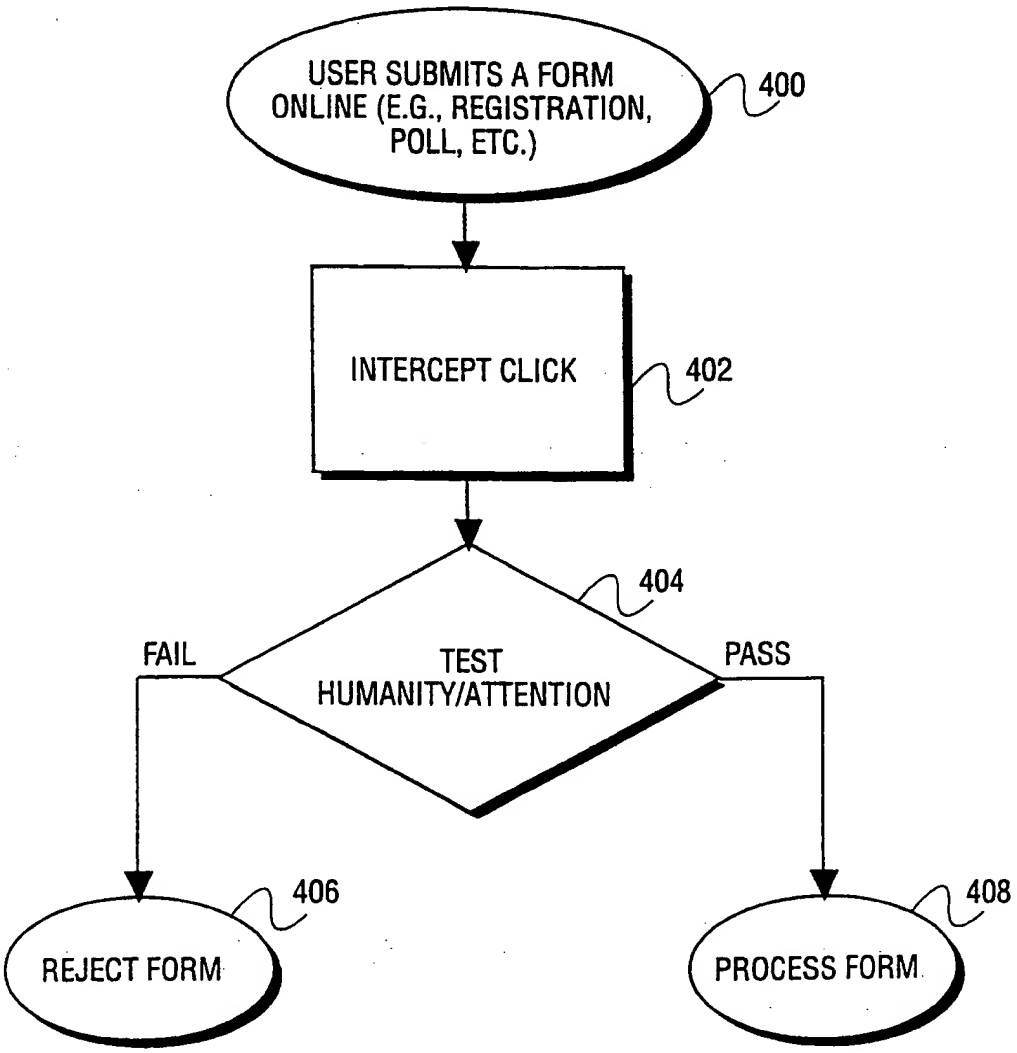

USER SUBMITS A FORM ONLINE (E.G., REGISTRATION, POLL, ETC.)
400
INTERCEPT CLICK
402
404
TEST HUMANITY/ATTENTION
FAIL
PASS
406
REJECT FORM
408
PROCESS FORM

FIG. 4

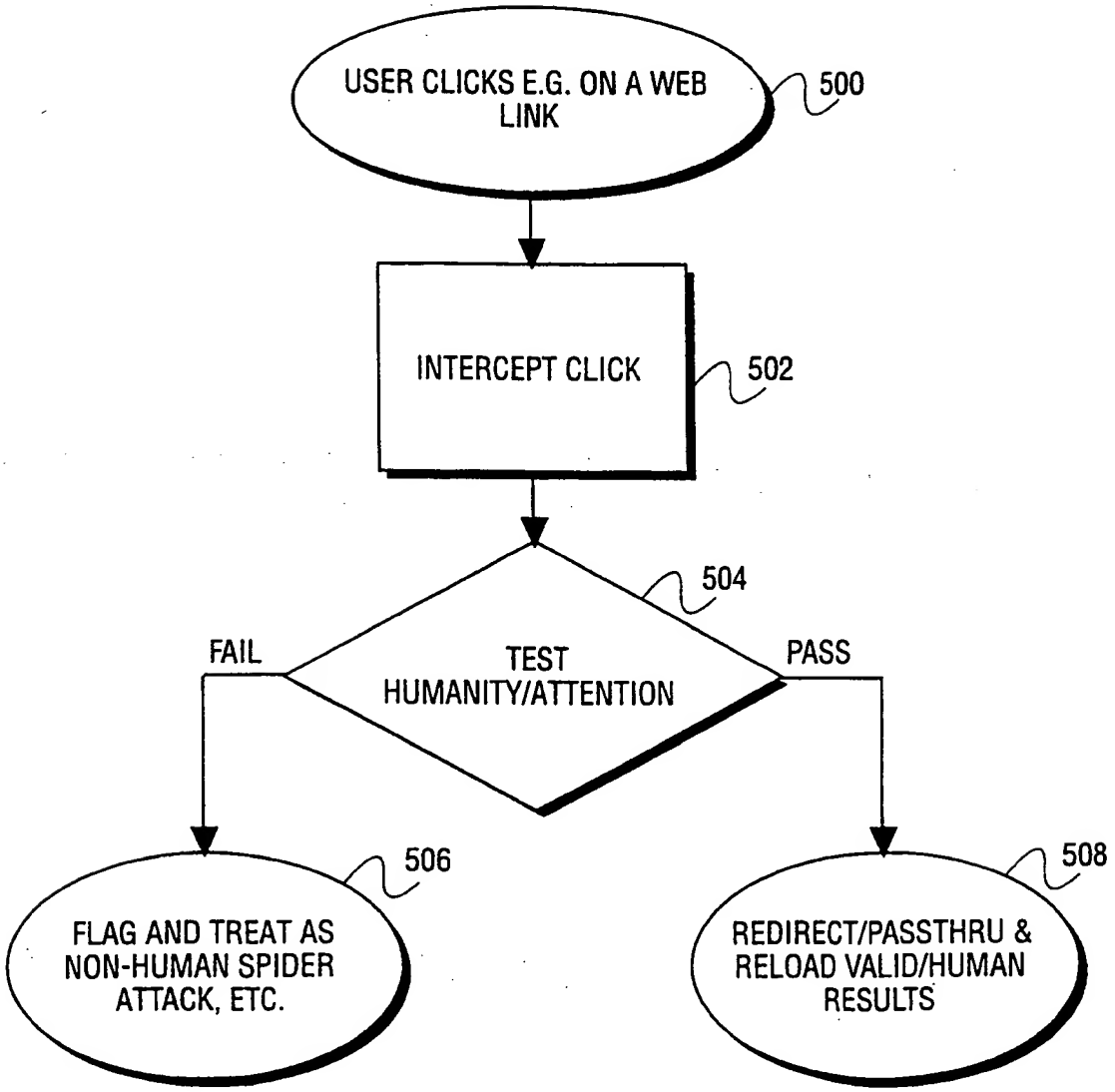

USER CLICKS E.G. ON A WEB LINK
500
INTERCEPT CLICK
502
504
TEST HUMANITY/ATTENTION
FAIL
PASS
506
FLAG AND TREAT AS NON-HUMAN SPIDER ATTACK, ETC.
508
REDIRECT/PASSTHRU & RELOAD VALID/HUMAN RESULTS

FIG. 5

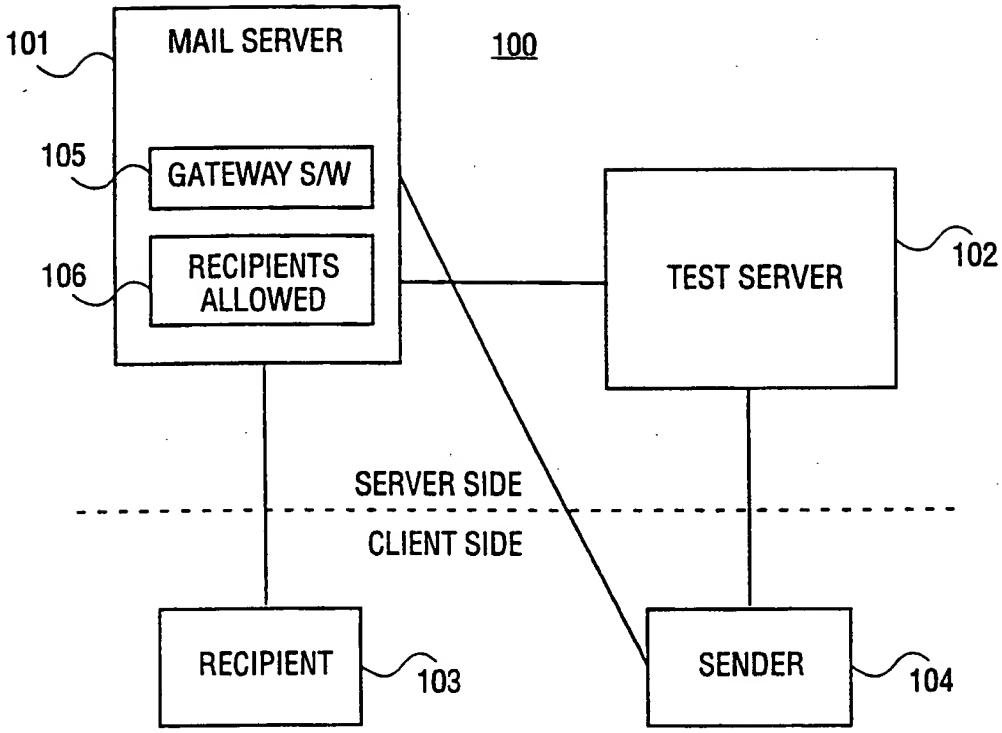

101
MAIL SERVER
100
105
GATEWAY S/W
106
RECIPIENTS ALLOWED
TEST SERVER
102
SERVER SIDE
CLIENT SIDE
RECIPIENT
103
SENDER
104

FIG. 6

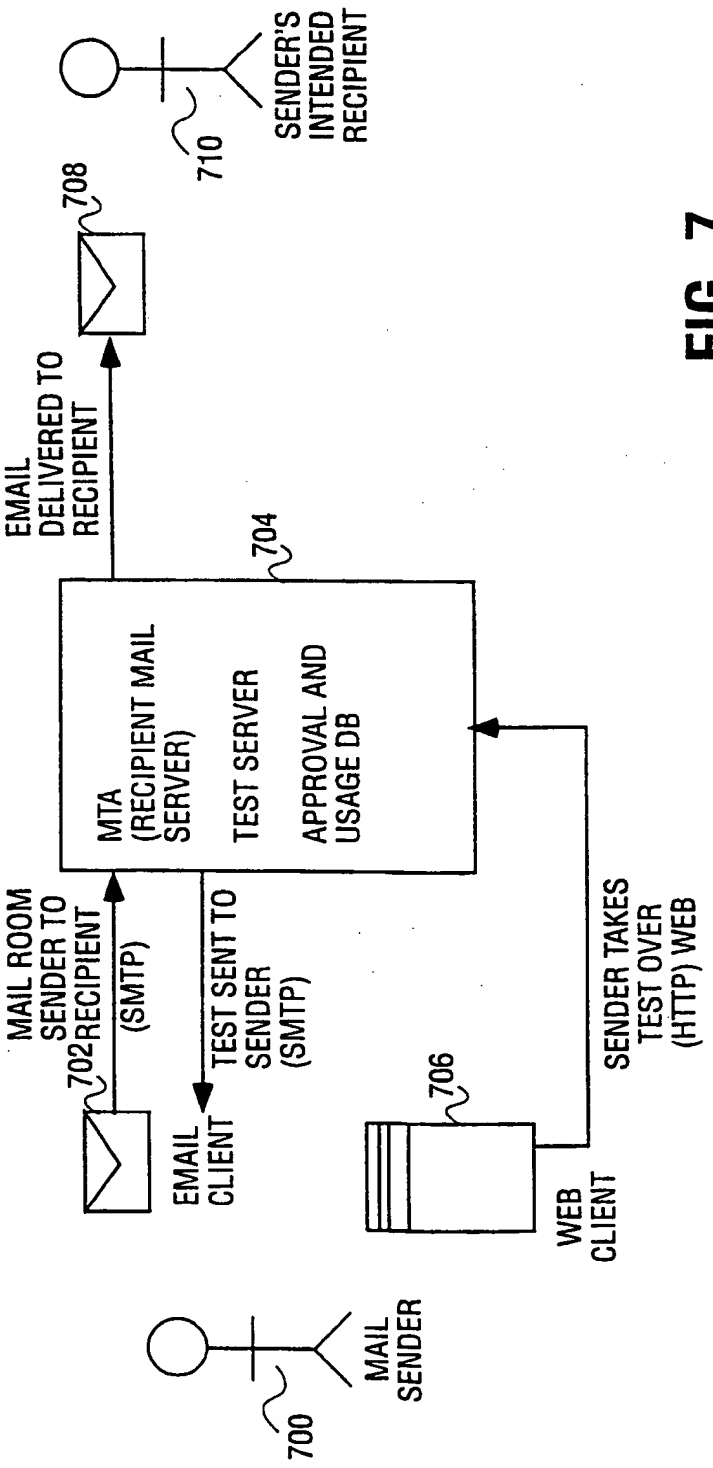
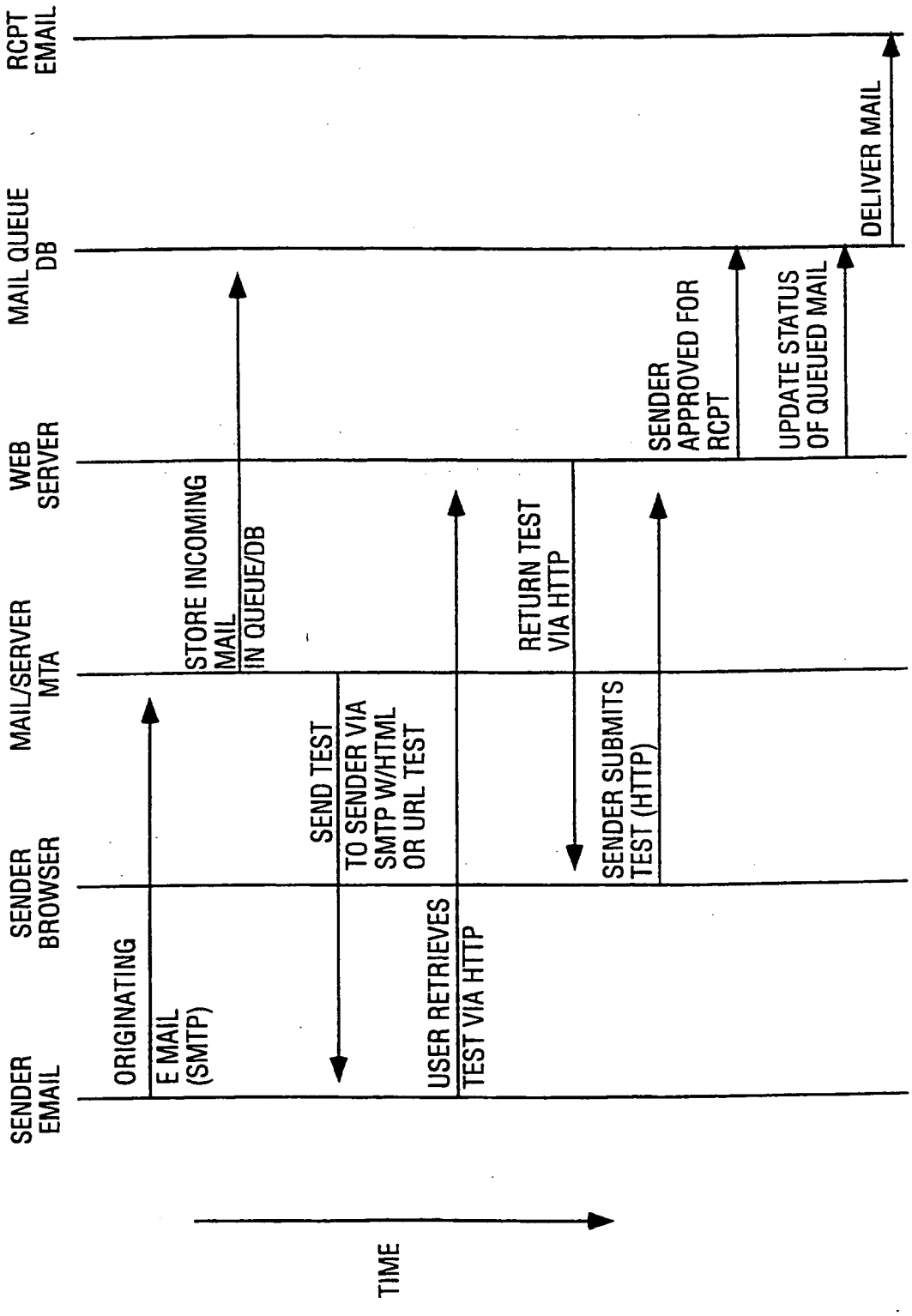

MAIL SENDER
700
EMAIL CLIENT
MAIL ROOM SENDER TO RECIPIENT (SMTP)
702
TEST SENT TO SENDER (SMTP)
MTA (RECIPIENT MAIL SERVER)
TEST SERVER
APPROVAL AND USAGE DB
704
EMAIL DELIVERED TO RECIPIENT
708
SENDER'S INTENDED RECIPIENT
710
WEB CLIENT
706
SENDER TAKES TEST OVER (HTTP) WEB

FIG. 7

FIG. 8


SENDER EMAIL
SENDER BROWSER
MAIL/SERVER MTA
WEB SERVER
MAIL QUEUE DB
RCPT EMAIL
ORIGINATING E MAIL (SMTP)
STORE INCOMING MAIL IN QUEUE/DB
SEND TEST TO SENDER VIA SMTP W/HTML OR URL TEST
USER RETRIEVES TEST VIA HTTP
RETURN TEST VIA HTTP
SENDER SUBMITS TEST (HTTP)
SENDER APPROVED FOR RCPT
UPDATE STATUS OF QUEUED MAIL
DELIVER MAIL
TIME

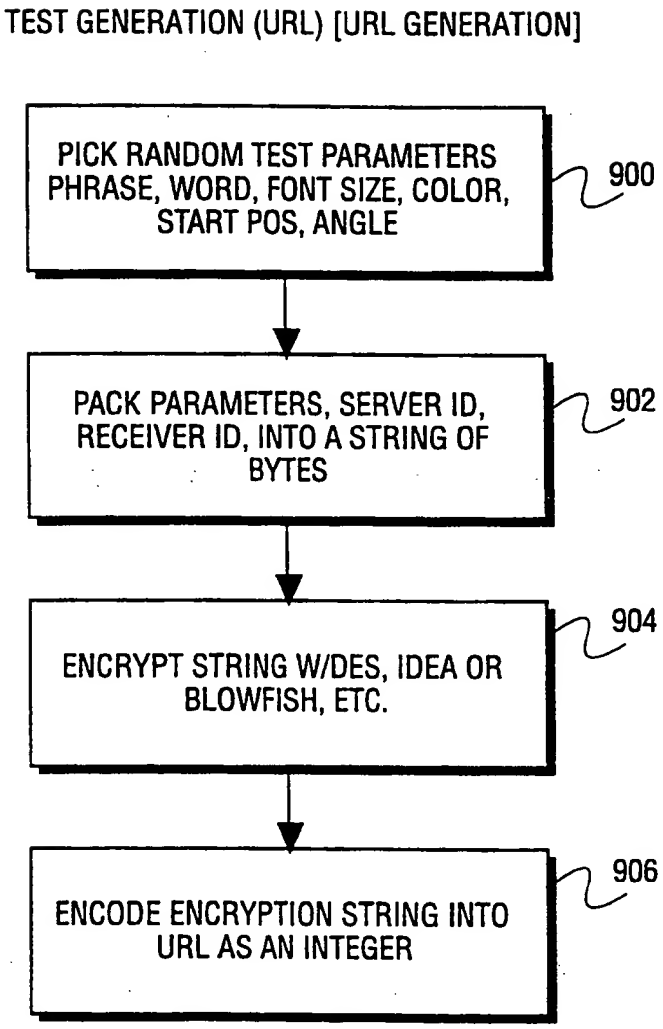

TEST GENERATION (URL) [URL GENERATION]
PICK RANDOM TEST PARAMETERS PHRASE, WORD, FONT SIZE, COLOR, START POS, ANGLE
900
PACK PARAMETERS, SERVER ID, RECEIVER ID, INTO A STRING OF BYTES
902
ENCRYPT STRING W/DES, IDEA OR BLOWFISH, ETC.
904
ENCODE ENCRYPTION STRING INTO URL AS AN INTEGER
906

FIG. 9

TEST GENERATION (URL DECODING, PRESENT TESTING)

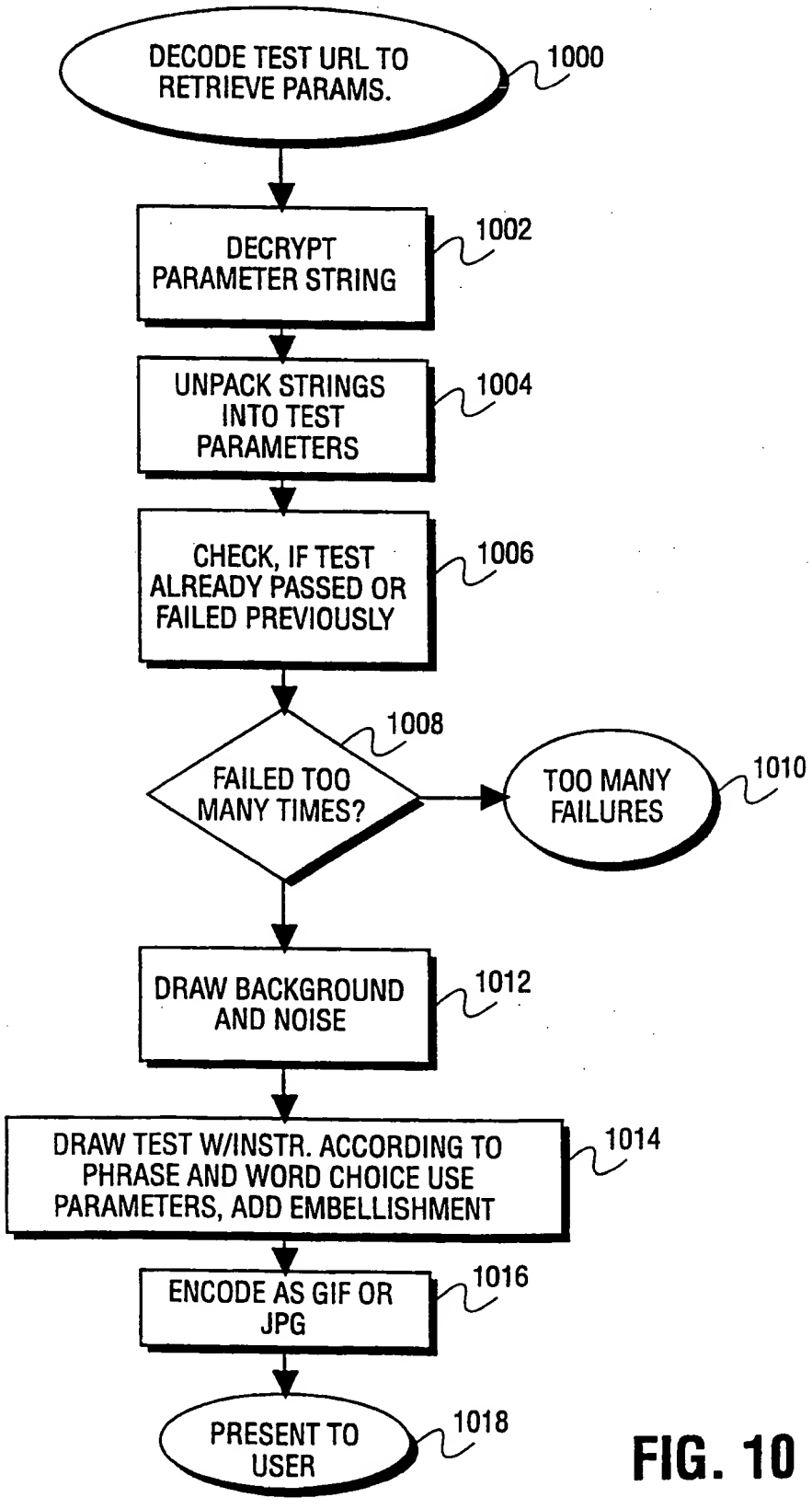


FIG. 10

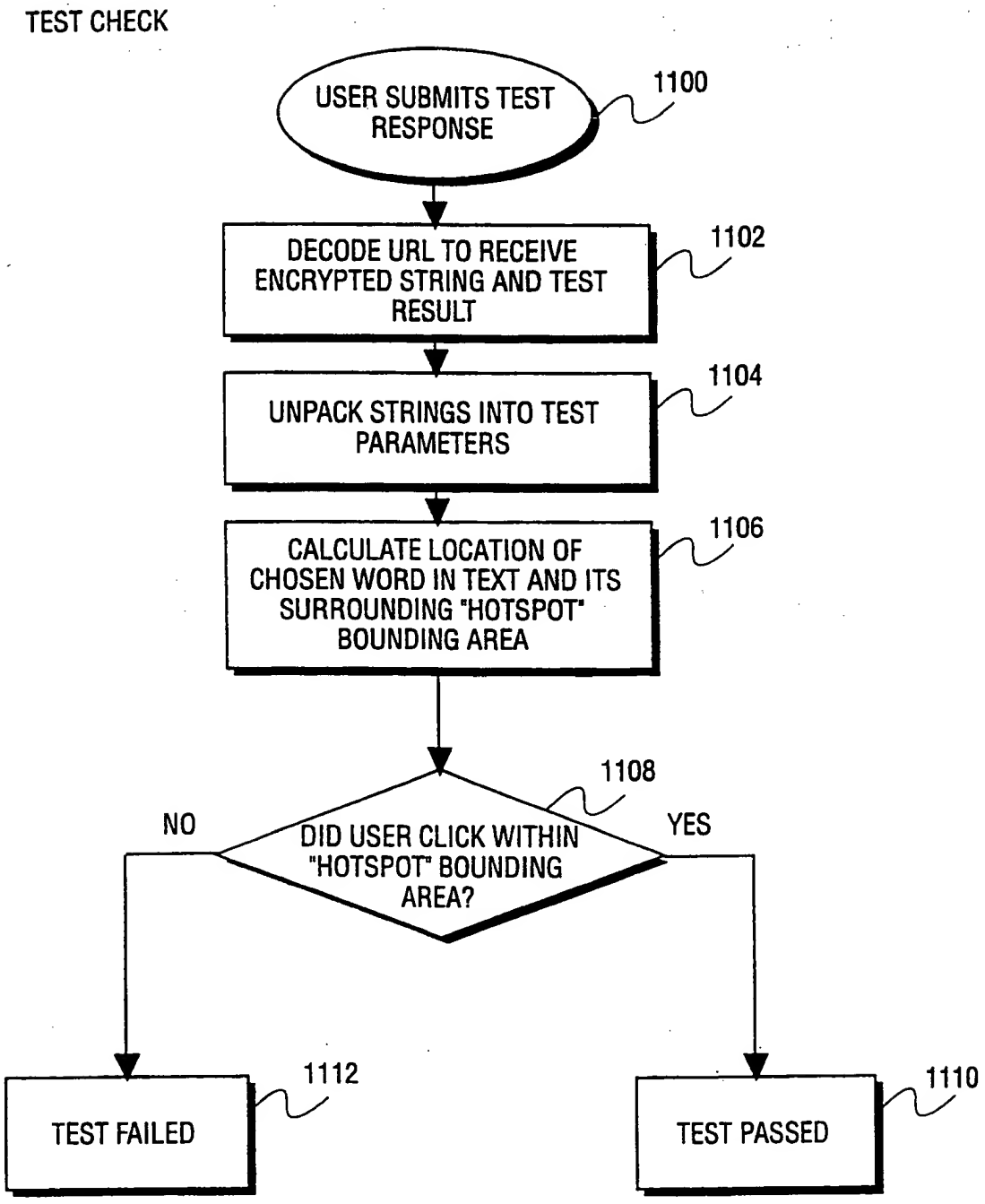

TEST CHECK
USER SUBMITS TEST RESPONSE
1100
DECODE URL TO RECEIVE ENCRYPTED STRING AND TEST RESULT
1102
UNPACK STRINGS INTO TEST PARAMETERS
1104
CALCULATE LOCATION OF CHOSEN WORD IN TEXT AND ITS SURROUNDING "HOTSPOT" BOUNDING AREA
1106
DID USER CLICK WITHIN "HOTSPOT" BOUNDING AREA?
1108
NO
YES
TEST FAILED
1112
TEST PASSED
1110

FIG. 11

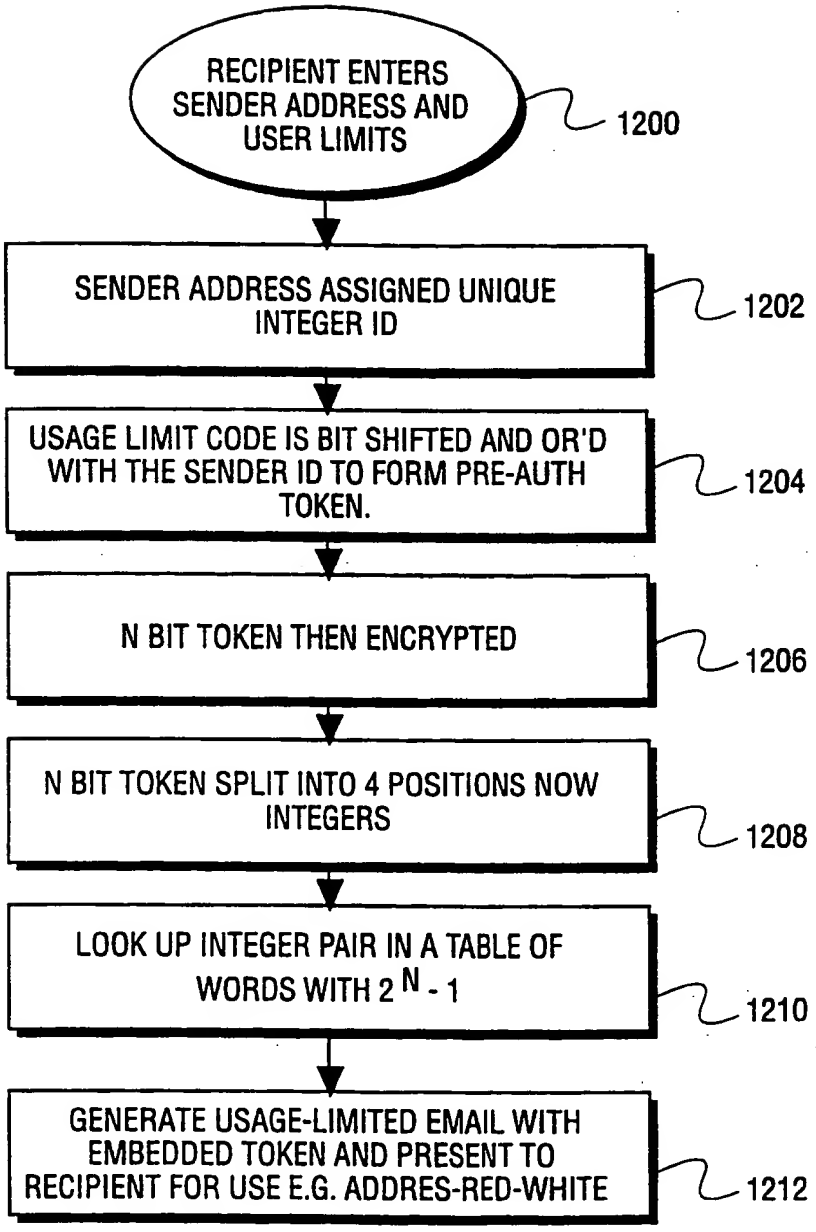

RECIPIENT ENTERS SENDER ADDRESS AND USER LIMITS
1200
SENDER ADDRESS ASSIGNED UNIQUE INTEGER ID
1202
USAGE LIMIT CODE IS BIT SHIFTED AND OR'D WITH THE SENDER ID TO FORM PRE-AUTH TOKEN.
1204
N BIT TOKEN THEN ENCRYPTED
1206
N BIT TOKEN SPLIT INTO 4 POSITIONS NOW INTEGERS
1208
LOOK UP INTEGER PAIR IN A TABLE OF WORDS WITH 2 N - 1
1210
GENERATE USAGE-LIMITED EMAIL WITH EMBEDDED TOKEN AND PRESENT TO RECIPIENT FOR USE E.G. ADDRES-RED-WHITE
1212

FIG. 12

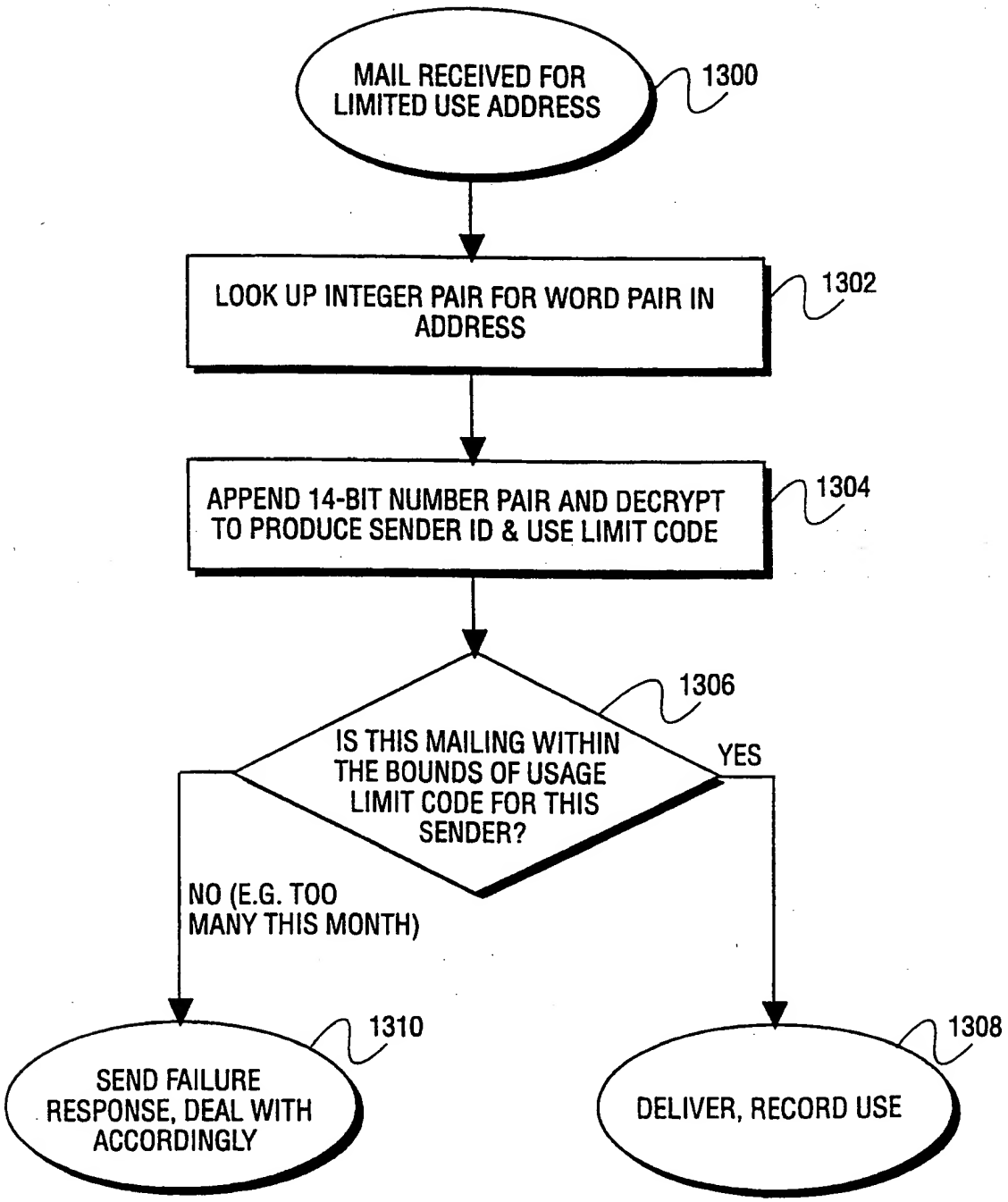

MAIL RECEIVED FOR LIMITED USE ADDRESS
1300
LOOK UP INTEGER PAIR FOR WORD PAIR IN ADDRESS
1302
APPEND 14-BIT NUMBER PAIR AND DECRYPT TO PRODUCE SENDER ID & USE LIMIT CODE
1304
IS THIS MAILING WITHIN THE BOUNDS OF USAGE LIMIT CODE FOR THIS SENDER?
1306
YES
NO (E.G. TOO MANY THIS MONTH)
SEND FAILURE RESPONSE, DEAL WITH ACCORDINGLY
1310
DELIVER, RECORD USE
1308

FIG. 13